(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 1 760 577 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
11.03.2009 Bulletin 2009/11

(51) Int Cl.:
G06F 3/06 *(2006.01)*

(21) Application number: **06022857.4**

(22) Date of filing: **05.09.2003**

(54) **Virtualization system and data access control method**

Virtualisierungssystem und Datenzugangskontrollverfahren

Système de virtualisation et procédé de contrôle d'accès à des données

(84) Designated Contracting States:
DE FR GB

(30) Priority: **18.09.2002 JP 2002271595**

(43) Date of publication of application:
07.03.2007 Bulletin 2007/10

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
03020209.7 / 1 400 893

(73) Proprietor: **HITACHI, LTD.**
Chiyoda-ku
Tokyo 100-8280 (JP)

(72) Inventors:
- **Innan, Masataka,**
  Hitachi,Ltd. Int. Prop. Office
  Chiyoda-ku,
  Tokyo 100-8220 (JP)
- **Shimada, Akinobu,**
  Hitachi,Ltd. Int. Prop. Office
  Chiyoda-ku,
  Tokyo 100-8220 (JP)
- **Tabuchi, Hideo,**
  Hitachi,Ltd. Int. Prop. Office
  Chiyoda-ku,
  Tokyo 100-8220 (JP)
- **Nakano, Toshio,**
  Hitachi,Ltd. Int. Prop. Office
  Chiyoda-ku,
  Tokyo 100-8220 (JP)

(74) Representative: **Strehl Schübel-Hopf & Partner**
Maximilianstrasse 54
80538 München (DE)

(56) References cited:
EP-A- 0 670 551
US-A1- 2002 087 751

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

EP 1 760 577 B1

Printed by Jouve, 75001 PARIS (FR)

Description

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] The present invention relates to a method for controlling a storage system, a storage controller and a storage system.

Description of the Related Art

[0002] The environment surrounding information systems is rapidly changing such as the progress of IT technology and shifting to broadband. Along with this, measures against the rapid increase of the amount of data handled are seriously required. Therefore, the storage systems operating in data centers are being subjected to prompt shifting to larger capacities and higher performances of the storage controllers such as disk array apparatuses.

[0003] To increase the storage area, a disk array apparatus can additionally be provided for example. However, accordingly as the number of the disk array apparatuses installed increases the load of management on the host computer side is inevitably increased, which may affect running of applications on the host computer side. Furthermore, recently, so-called "virtualization" is in progress that is a technology for managing a huge storage area efficiently without making users aware of the hardware configuration as far as possible. However, the virtualization may cause an increase of the load of processing on the host computer side.

[0004] In view of such circumstances, it is expected that the need for the technology for reducing the load of processing on the host computer side will become greater and greater. A technology is thus being sought that ensures larger capacities and higher performances of storage controllers without increasing the load of processing on the host computer side.

[0005] US 2002/0087751 A1 discloses a storage system in which a data switch controlled by a storage controller is coupled between a host computer and a plurality of storage devices. The data switch includes a memory for storing data routing information used for routing data directly between the host computer and the storage devices without passing through the storage controller.

SUMMARY OF THE INVENTION

[0006] The present invention was conceived on the basis of the above background and one object thereof is to provide a method for controlling a storage system, a storage controller and a storage system so as to efficiently perform an access limit function and data copy control.

[0007] This object is achieved by the methods defined in claims 1 and 2 and the corresponding systems defined in claims 4 and 5.

[0008] Features and objects of the present invention other than the above will become clear by reading the description of the present specification with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] For a more complete understanding of the present invention and the advantages thereof, reference is nowmade to the following description taken in conjunction with the accompanying drawings wherein:

Fig. 1 illustrates the configuration of a storage system according to one embodiment of the present invention;
Fig. 2 illustrates the structure of a data frame on a Fiber Channel according to one embodiment of the present invention;
Fig. 3 illustrates the structure of a frame header according to one embodiment of the present invention;
Fig. 4 illustrates the structure of FCP_CMND according to one embodiment of the present invention;
Fig. 5 is a flow chart illustrating the operation of a storage system according to one embodiment of the present invention when a data frame is transmitted from a host computer to a first storage controller;
Fig. 6 shows an LUN management table according to one embodiment of the present invention;
Fig. 7 shows an access limit management table according to one embodiment of the present invention;
Fig. 8 is a flowchart illustrating an access limit function according to one embodiment of the present invention;
Fig. 9 shows a priority management table according to one embodiment of the present invention;
Fig. 10 is a flow chart illustrating a priority control function according to one embodiment of the present invention;
Fig. 11 shows a pair management table according to one embodiment of the present invention; and
Fig. 12 is a flow chart illustrating the processing of the storage system according to one embodiment of the present invention when a data frame is transmitted in which a data writing request for a main logic volume is described.

DETAILED DESCRIPTION OF THE INVENTION

[0010] At least the following matters will be made clear by the explanation in the present specification and the description of the accompanying drawings.

[Summary of the Disclosure]

[0011] A method for controlling a storage system, the storage system comprises a host computer; a first storage controller connected communicably to the host computer, for receiving a data frame transmitted from the host

computer and executing data input to and data output from a first storage device in response to a data input/output request described in the data frame; and a second storage controller connected communicably to the first storage controller, for receiving a data frame transmitted from the host computer to the first storage controller and relayed by the first storage controller and executing data input to and data output from a second storage device in response to a data input/output request described in the data frame, wherein when the first storage controller receives the data frame transmitted from the host computer the first storage controller relays the data frame to the second storage controller in response to information described in the data frame.

[0012] In this method, the second storage controller is connected to the first storage controller such that the data frame transmitted from the host computer is received temporarily by the first storage controller and is transmitted by the first storage controller to the second storage controller in response to information described in the data frame. In this configuration, the host computer transmits the data frame to the first storage controller indiscriminately regardless of whether the data frame that the host computer is going to transmit is directed to the first storage controller or to the second storage controller. Therefore, the host computer has no need to communicate directly with the second storage controller when transmitting the data frame and the load of processing in the host computer can be reduced by the amount to be produced by the direct communication, compared to the scheme in which a host computer executes direct communications with the second storage controller.

[0013] A function for relaying the data frame to the second storage controller in response to information described in the data frame when the data frame transmitted from the host computer has been received can be realized by, for example, a program that is run on the first storage controller.

[0014] The data frame is, for example, a data frame conforming to Fiber Channel Protocol. The relaying is executed by a Fiber Channel switch included in the first storage controller.

[0015] The information is at least one of information for specifying a Fiber Channel port of the source of the data frame, information for specifying a Fiber Channel port of the destination of the data frame and information for specifying a storage device. Among these, the information for specifying a Fiber Channel port of the source is, for example, Source_ID described later. The information for specifying a Fiber Channel port of the destination is, for example, Destination ID described later. The information for specifying a storage device is, for example, FCP_LUN described later.

[0016] The storage device is, for example, a logic volume organized in zones on a storage area provided by a disk drive or a physical storage area provided by a disk drive.

[0017] In this method, the first storage controller functions rather as a network switch which either receives by itself the data frame transmitted from the host computer or relays the data frame to the second storage controller. Then, because the first storage controller functions as a network switch, an applied operation can be easily provided in which the information and the functions possessed by the first storage controller are combined with the function as a network switch. Therefore, functions such as, for example, access limit function, priority control, double management of data, described later, can be easily provided. Not limited to these functions, various functions can be easily realized which are obtained by combining the function as the network switch and the functions that the first storage controller originally has as a storage controller.

[0018] Furthermore, when the virtualization technology is applied to the storage system, a plurality of storage devices are managed virtually unitarily as one device and/or management/operation of functions such as remote copy and data copy management are consolidated. By being provided with such functions, influence on the resource consumption and the processing rate on the host computer side becomes inevitable. However, by making available the provision of the functions formed by combining the functions that the first storage controller originally has as a storage controller and the function as that of a network switch by providing the first storage controller with the function as that of a network switch as described above, it is possible to provide the functions that have been originally included in the host computer, to the side of the first storage controller, to minimize the resource consumption and the load of processing on the host computer side and to provide a system supporting the virtualization.

[0019] Furthermore, when a general-purpose switch such as a Fiber Channel switch is used as the network switch, the first storage controller and the second storage controller can be easily connected even though those storage controllers are different models or are from different manufacturers. Therefore, the storage system can be easily configured and full use of resources such as existing or obsolete storage controllers can be progressed. In addition, in embodiments of (i) basic operations, (ii) access limit function, (iii) priority control and (iv) data copy management, full use of a plurality of storage controllers can be progressed by connecting one storage controller provided with these functions to a plurality of storage controllers not provided with these functions.

[0020] In the data frame, as the information, at least one of information for specifying a Fiber Channel port of the source of the data frame, information for specifying a Fiber Channel port of the destination of the data frame and information for specifying a storage device, is described. The first storage controller stores access limit information indicating whether an access of a Fiber Channel port of the source to a Fiber Channel port or a storage device of the destination is permitted or not permitted. Therefore, it can be arranged that, when the first

storage controller receives the data frame from the host computer, the first storage controller refers to the access limit information to check whether an access of a Fiber Channel port of the source of the data frame to a Fiber Channel port of the destination or a storage device of the destination is permitted and relays the data frame to the second storage controller only when the access is permitted.

[0021] In this manner, the first storage controller judges based on the access limit information whether the first storage controller should transmit the data frame to the second storage controller and, thus, the first storage controller executes replacing the role of the second storage controller the process relating to the access limit of the data frame directed to the second storage controller. Therefore, it is enough that the function for processing an access limit is implemented only on the first storage controller, and the load of maintenance is reduced by the amount to be created by taking the trouble of providing such a mechanism to the second storage controller. The function of the first storage controller relating to the access limit is realized by, for example, a program that is run on the first storage controller.

[0022] Since there is no need for the second storage controller to be provided with access limit function, a storage controller having lower performance than that of the first storage controller can be used as the second storage controller and, therefore, the introduction cost and the operation cost can be reduced. Furthermore, old types of storage controllers having lower performance compared to the newly introduced storage controllers can be fully used as second storage controllers. Yet furthermore, it is enough that the access limit management table is stored only in the first storage controller and, therefore, reduction of the load of maintenance created by the unified management is progressed.

[0023] Arrangement may be as follows: in the data frame, as the information, information for specifying a Fiber Channel port of the source of the data frame and information for specifying a Fiber Channel port of the destination of the data frame are described; the first storage controller stores information indicating the priority of data input/output process corresponding to the data frame by relating to combinations of a Fiber Channel port of the source and a Fiber Channel port of the destination; when the first storage controller receives the data frame from the host computer, the apparatus executes data input to and data output from the storage device connected to a port of the destination of the data frame according to the information indicating the priority. The information indicating the priority is, for example, information indicating the timing for executing the data input to and data output from the storage device and, more specifically, a delay time which will be described later.

[0024] In this mechanism, it can be considered that the first storage controller executes the priority control function replacing the role of the second storage controller. That is, by only implementing the priority control function on the first storage controller, a priority control can also be executed to the data frame directed to the second storage controller and the load of operation imposed on the second storage controller can be reduced by the amount to be created by providing such a mechanism.

[0025] The function of the first storage controller relating to the priority control is realized by, for example, a program that is run on the first storage controller.

[0026] Because there is no need for the second storage controller to be provided with priority control function a storage controller having lower processing performance than that of the first storage controller can be used as the second storage controller and, therefore, the introduction cost and the operation cost can be reduced. Furthermore, old types of storage controllers having lower performance compared to the newly introduced storage controllers can be fully used as second storage controllers. Yet furthermore, it is enough that the priority management table is stored only in the first storage controller, thus achieving a reduction of the load of maintenance by the unified management.

[0027] A method for controlling a storage system, the storage system comprising a host computer; a first storage controller connected communicably to the host computer, for receiving a data frame transmitted from the host computer and executing data input to and data output from a first storage device in response to a data input/output request described in the data frame; and a second storage controller connected communicably to the first storage controller, for receiving a data frame transmitted from the host computer to the first storage controller and relayed by the first storage controller and executing data input to and data output from a second storage device in response to a data input/output request described in the data frame, wherein the first storage controller stores a copy of data stored in the first storage device which is the object of data input/output effected by the first storage controller, into the second storage device as well which is the object of data input/output effected by the second storage controller, the data frame having therein described information for specifying a Fiber Channel port connected to the first storage device or the second storage device which is the destination of the data frame and information for specifying the storage devices, wherein when the first storage controller receives the data frame in which is described a data writing request for the first storage device the first storage controller executes writing of data on that data frame to the first storage device while simultaneously the first storage controller creates a data frame in which the information for specifying the Fiber Channel port and the information for specifying the storage devices described in the data frame are rewritten such that the object is the second storage device, the first storage controller transmitting the created data frame to the second storage controller, and wherein the second storage controller receives the data frame and executes writing of data on that data frame to the second storage device, to thereby store a copy of the data stored in the

first storage device into the second storage device as well that is the object of data input/output effected by the second storage controller.

[0028] According to this method, a mechanism for data copy management in which data stored in the first storage device is stored into the second storage device as well can be easily realized only by providing a mechanism for creating a data frame as described above in the first storage controller.

[0029] It is to be noted that the function of the first storage controller relating to the data copy management is realized by, for example, a program that is run on the first storage controller.

[The Configuration of the Storage System]

[0030] The configuration of a storage apparatus system (storage system) described as an embodiment of the present invention is illustrated in Fig. 1. The storage system is operated in, for example, a data center. A host computer 5 and a first storage controller 10 are connected through first communication means 40 conforming to Fiber Channel Protocol. The host computer 5 is a computer such as, for example, a personal computer, a workstation and a general-purpose computer. The first storage controller 10 and a second storage controller 20 are connected through second communication means 50 conforming to the Fiber Channel Protocol.

[0031] The first storage controller 10 includes a switching control unit 11 and a storage device control unit 12. The switching control unit 11 and the storage device control unit 12 are connected by an internal communication path 13 such as an internal bus.

[0032] The switching control unit 11 comprises a switching device 111, a microprocessor 112, and a control memory 113 such as a ROM, a RAM, etc. The switching device 111 is provided with functions as a Fiber Channel switch (hereinafter referred to as "FC switch") and has a predetermined number of Fiber Channel ports. The switching device 111 receives a data frame transmitted from the host computer 5 and relays the data frame to the storage device control unit 12 or the second storage controller 20. In addition, the switching device 111 also receives a data frame transmitted from the storage device control unit 12 or the second storage controller 20 and relays the data frame to the host computer 5. The microprocessor 112 provides various functions of the switching control unit 11. In addition, the microprocessor 112 controls various components included in the switching control unit 11. Furthermore, the microprocessor 112 controls transmitting/receiving of the data frame conducted at the switching device 111. The control memory 113 is utilized as, for example, a work memory for processes executed by the microprocessor 112, or a storage area for various data such as tables and variables.

[0033] The storage device control unit 12 receives a data frame transmitted through the internal communication path 13 and executes data input to and data output from a disk drive 14 in response to a data input/output request described in the data frame. The data input/output processes include a data writing request and a data read-out request. In the following description, data input and output includes the cases where it includes either a data writing request or a data read-out request and where it includes both of a data writing request and a data read-out request.

[0034] The microprocessor 121 provides various functions of the storage device control unit 12, and controls the various components included in the storage device control unit 12, etc. A control memory 123 is utilized as, for example, a work memory for the processes executed by the microprocessor 121 and as a storage area for various data such as tables and variables. A cache memory 124 is used as, for example, temporary storage means for data being input to/output from the disk drive 14. A drive controller 125 executes data input/output processes to the disk drive 14. The drive controller 125 may operate the disk drive 14 in RAID (Redundant Array of Inexpensive Disks) scheme. The drive controller 125 manages a physical storage area provided by the disk drive 14 by dividing it into logic volumes that are logic storage areas organized on the physical storage area. A specific identifier called LUN (Logical Unit Number) is given to each logic volume respectively. When an LUN has been designated, one ormore disk drives 14 are identified. Furthermore, at least one or more logic volumes are related to one Fiber Channel port and, by designating one Fiber Channel port, all the logic volumes related to it are identified. The disk drive 14 can be held in the same housing as that of the first storage controller 10 or in housing different from that of the first storage controller 10.

[0035] The second storage controller 20 comprises a switching device 201, a microprocessor 202, a control memory 203, a cache memory 204, a drive controller 205 and a disk drive 21. The switching device 201 is provided with functions as a Fiber Channel switch (hereinafter referred to as "FC switch") having a plurality of Fiber Channel ports and second communication means 50 is connected to the Fiber Channel ports.

[0036] The microprocessor 202 provides various functions of the second storage controller 20 and controls various components included in the second storage controller 20, etc. The control memory 203 is utilized as, for example, a work memory for the processes executed by the microprocessor 202, or a storage area for various data such as tables and variables. The cache memory 204 is used as, for example, temporary storage means for data input to/output from the disk drive 21. The drive controller 205 executes data input/output processes to the disk drive 21.

[0037] The drive controller 205 may operate the disk drive 21 inRAID (Redundant Array of Inexpensive Disks) scheme. Thedrive controller 205 manages a physical storage area provided by the disk drive 21 by dividing it into logic volumes that are logic storage areas organized

on the physical storage area. A specific identifier called LUN (Logical Unit Number) is given to each logic volume respectively. When an LUN has been designated, one or more disk drives 21 are identified.

[0038] The disk drive 21 can be held in the same housing as that of the second storage controller 20 or in housing different from that of the second storage controller 20.

[Data Frame]

[0039] Next, it will be described the structure of data frames conforming to the Fiber Channel Protocol, that is transmitted and received between the host computer 5, the first storage controller 10 and the second storage controller 20. The details of the Fiber Channel Protocol and data frames conforming to it are described in, for example, "Technical Instruction on Fiber Channel" edited by the Fiber Channel Technical Association, Ronso-sha, etc.

[0040] The structure of a data frame 70 on a Fiber Channel is illustrated in Fig. 2. At the head of the data frame 70, an SOF (Start of Frame) (71) of four bytes is placed and, following the SOF, a frame header (72) of 24 bytes, a data field (73) of 2112 bytes and a CRC (Cyclic Redundancy Check) (74) of four bytes are provided. At the tail of the data frame, an EOF (75) of four bytes is placed.

[0041] The structure of the frame header (72) of 24 bytes shown in Fig. 2 is illustrates in Fig. 3. Destination_ID (81) is an identifier of the Fiber Channel port being the destination of the data frame and N_Port_ID, N_Port_Name or etc. is described in it. Source_ID (82) is an identifier of the Fiber Channel port of the source of the data frame and N_Port_ID, N_Port_Name or etc. is described in it.

[0042] The structure of FCP_CMND (Fiber Channel Protocol for SCSI Command) being one of the payloads of the data field shown in Fig. 2 is illustrated in Fig. 4. In FCP_LUN (FCP LogicaLUNit Number) (91), a LUN of the logic volume to which a SCSI command described in the payload is to be applied. In FCP_CNTL (FCP Control) (92), a command control parameter is described. In FCP_CDB (FCP Command Descriptor Block) (93), a SCSI command CDB (Command Data Block) is described. In FCP_DL (FCP Data Length) (94), the upper limit of the data size that can be converted between an initiator and a target is described.

[Basic Operation]

[0043] Next, the operation of the storage system while a data frame is transmitted from the host computer 5 to the first storage controller 10 will be described referring to a flow chart shown in Fig. 5.

[0044] First, a data frame is transmitted from the host computer 5 to a Fiber Channel port of the switching device 111 of the first storage controller 10 through the first communication means 40 (S511). When the switching device 111 of the first storage controller 10 receives the data frame (S512), the switching device 111 transmits an interrupt request to the microprocessor 112 (S513). When the microprocessor 112 receives the interrupt request, the microprocessor 112 stores in the control memory 113 the contents of Destination_ID (81), Source_ID (82) and FCP_LUN (91) of the received data frame (S514).

[0045] In the control memory 113 of the first storage controller 10, a LUN management table shown in Fig. 6, in which whether the FCP_LUN (91) is directed to a logic volume of the first storage controller 10 or of the second storage controller 20 is described is stored. The microprocessor 112 checks whether the data frame is directed to a logic volume of the first storage controller or of the second storage controller by referring the FCP_LUN (91) of the data frame stored in the control memory 113 to the LUN management table (S615).

[0046] Now, if the data frame is directed to a logic volume of the first storage controller 10, the first storage controller 10 transmits the data frame to the storage device control unit 12 through the internal communication path 13 (S516). When the storage device control unit 12 receives the data frame, the storage device control unit 12 executes data input and output corresponding to the SCSI command described in the data frame to the logic volume identified from the contents of FCP_LUN (91) of the data frame (S517). After executing the data input/output process, the first storage controller 10 transmits to the switching control unit 11 the data frame in which the result of the process and data read out from the logic volume etc. are described, when necessary (S518). Then, when the switching device 111 receives the data frame, it transmits the data frame to the host computer 5 (S519).

[0047] On the other hand, in (S515), if the data frame received from the host computer 5 is directed to a logic volume of the second storage controller 20, the microprocessor 112 controls the switching device 111 such that the switching device 111 relays the data frame to the second storage controller 20 (S520). When the second storage controller 20 receives the data frame (S521), it executes data input and output corresponding to the SCSI command described in the data frame to a logic volume identified from the contents of FCP LUN (91) of the data frame (S522).

[0048] After executing the data input/output process, the second storage controller 20 transmits to the first storage controller 10 through the second communication means the data frame in which the report of completion, the result of the process or data read out from the logic volume are described, when necessary (S523). In the Destination_ID (81) of the data frame, information for specifying a Fiber Channel port of the host computer 5 is described. When the switching device 111 of the first storage controller 10 receives the data frame (S524), it relays the data frame to the host computer 5 (S525).

[0049] As described above, the data frame transmitted

from the host computer 5 and input into the switching device 111 of the first storage controller 10, is judged at the switching control unit 11 whether it is directed to the first storage controller 10 or the second storage controller 20 and, if it is directed to the second storage controller 20, the data frame is relayed from the switching device 111 to the second storage controller 20. In other words, the first storage controller 10 receives replacing the role of the second storage controller 20 the data frame that the host computer 5 originally has directed to the second storage controller 20.

[0050] Furthermore, by receiving the data frame by the first storage controller 10 replacing the role of the second storage controller 20, the host computer 5 only has to communicate with the first storage controller 10 and the load of processing of the host computer 5 can be reduced by the amount to be created by the communication between the host computer 5 and the second storage controller 20.

[0051] Yet furthermore, by managing the LUN management tables unifying at the first storage controller 10, the logic volumes of the first storage controller 10 and the second storage controller 20 can be managed being unified and, therefore, the load of maintenance imposed on work such as updating of the LUN management tables by operators etc. can be reduced. Especially when the disk drive 14 and disk drive 21 are operated while being virtualized, such a unified management scheme considerably contributes to more efficient operation and reduction of the load of operation.

[0052] When a disk array apparatus is the one equipped with a very ordinary interface as the connection interface with the host computer 5, like the switching device 201 as described above, the disk array apparatus can be used as the second storage controller 20 of the embodiment and, for example, in a storage system using a disk array apparatus having a higher performance as the first storage controller 10 and a disk array apparatus having a lower performance than the other one as the second storage controller 20, a merit can be produced that the functions of the first storage controller 10 can be applied to the second storage controller 20.

[Access Limit Function]

[0053] Next, the access function included in the first storage controller 10 will be described. The first storage controller 10 stores the access limit management table exemplified in Fig. 7. In the access limit management table, combinations of Destination_ID, Source ID and LUN, that do not permit any data input/output process are described. The access limit function is a function that does not permit any data input/output process corresponding to a data input/output request described in the data frame when the combination of Destination_ID, Source_ID and LUN described in the data frame received by the first storage controller 10 from the host computer 5 is described in the access limit management table. The contents of the access limit management table may be registered automatically by a function of the first storage controller 10 or may be registered in a manual operation by an operator.

[0054] The access limit function will be described in detail referring to the flow chart shown in Fig. 8. This flow chart describes the case where a data frame in which a data input/output request directed to a logic volume of the second storage controller 20 is described is transmitted from the host computer 5 to the first storage controller 10.

[0055] When the first storage controller 10 receives a data frame (S812) transmitted from the host computer 5 (S811), the first storage controller 10 checks whether the combination of Destination_ID (81), Source_ID (82) and FCP_LUN (91) in the data frame is described in the access limit management table (S813, S814). If the combination is not described in the access limit management table, the first storage controller 10 relays the data frame to the second storage controller 20 (S815).

[0056] On the other hand, If the combination is described, the data frame is transmitted from the switching control unit 11 to the storage device control unit 12 (S816). When the storage device control unit 12 receives the data frame, the storage device control unit 12 refers to the access limit management table in the control memory 113 (S817). If the combination of Destination_ID (81), Source_ID (82) and FCP_LUN (91) in the data frame is a combination not to be permitted the access, the storage device control unit 12 creates a data frame directed to the host computer 5, in which it is described that it is not permitted to access, and transmits the newly created data frame to the switching device 111 (S818). Then, the newly created data frame is transmitted to the host computer 5 through the switching device 111.

[0057] On the other hand, if the received data frame is a data frame to be permitted the access, the storage device control unit 12 transmits the data frame to the switching device 111 (S819) and the switching device 111 relays the data frame to the second storage controller 20 (S820).

[0058] When the second storage controller 20 receives the data frame (S821), the second storage controller 20 executes data input and output corresponding to the SCSI command described in the FCP_CDB in the data frame to the logic volume described in the FCP_LUN in the data frame (S822). After executing the data input/output process, the second storage controller 20 transmits to the first storage controller 10 the data frame in which the result of the process and data read out from the logic volume etc. are described, when necessary (S823).

[0059] In this embodiment, combinations of Source_ID, Destination_ID and LUN that do not permit any data input/output process are described in the access limit management table, but instead, combinations of Source_ID, Destination_ID and LUN that permit data input/output processes may be described in the table. The access limit can be set in various forms including not only

the case where access limit is executed for all the combinations of Source_ID, Destination_ID and LUN but also the casewhere it is executed for any one of Source_ID, Destination_ID and LUN or for a combination of any two or more of these.

[0060] As described above, the first storage controller 10 executes replacing the role of the second storage controller 20 processes relating to the access limit. Therefore, it is enough that the function for executing the access limit is implemented on the first storage controller 10. Therefore, the load of maintenance relating to the access limit function can be reduced by the amount to be created by providing such a mechanism to the second storage controller 20. Furthermore, because no access limit function needs to be included in the second storage controller 20, a storage controller having a lower processing capacity than the first storage controller 10 can be used as the second storage controller 20 and, therefore, this is effective for reducing of the introduction cost and the operation cost.

[0061] Furthermore, old types of storage controllers having lower performance compared to the newly introduced storage controllers can be fully used as second storage controllers 20. Yet furthermore, it is enough that the access limit management table is stored only in the first storage controller 10 and, therefore, reduction of the load of maintenance created by the unified management is progressed.

[Priority Control]

[0062] Next, the priority control function included in the first storage controller 10 will be described. The first storage controller 10 stores the priority management table shown in Fig. 9. In the priority management table, the relations of Destination_ID, Source_ID and the delay time are described. The delay time is used as, for example, a parameter for determining the starting time of execution of a process that may increase the load of processing of the microprocessor 121. The priority control function is a function that controls the order of execution of the data input and output corresponding to the data frame received by the first storage controller 10 from the host computer 5, according to the delay time set in the priority management table. The contents of the priority management table may be registered automatically by a function of the first storage controller 10 or may be registered in a manual operation by an operator.

[0063] The priority control function will be described in detail referring to the flow chart shown in Fig. 10. When the first storage controller 10 receives a data frame A (S1012) transmitted from the host computer 5 (S1011), the first storage controller 10 checks whether the combination of Source_ID and Destination_ID in the data frame A exists in the priority management table or not (S1013, S1014).

[0064] If the combination exists, the first storage controller 10 determines the combination as the object for the access priority when a delay time is not set in the combination (S1015), executes data input and output for data frame A (S1016) and, after execution, transmits to the host computer 5 the data frame in which the report of completion is described (S1017).

[0065] If a delay time is set in (S1015), the first storage controller 10 checks whether another data frame B exists waiting to be processed in a queue managed by the control memory 123 in the storage device control unit 12 (S1020) and, if the data frame B exists waiting, causes the starting of execution of the process for data frame A stand by for the period of the delay time set in the priority management table corresponding to the contents of the data frame A (S1021). The starting time of which processes for the data frame should be delayed is adequately determined taking into consideration the load of processing imposed on the microprocessor 121 such that the data input/output process will be executed in an adequate order according to the needs of users. Arrangement may be such that the control of the order of executing processes using the delay time as a parameter can be determined automatically by the storage system or that the control of the order can be set by a user from a control terminal connected to the storage system. After the delay time has passed, the first storage controller 10 executes data input and output for data frame A (S1022) and transmits to the host computer 5 a data frame in which the report of completion is described (S1023).

[0066] On the other hand, in (S1020), if another data frame B does not exist waiting in the queue, the first storage controller 10 executes processes according to the ordinary procedure for data frame A described in the basic operation above. That is, if the data frame A is directed to the disk drive 14 in the first storage controller 10, data input/output process for the disk drive 14 is executed in the first storage controller 10 and, if the data frame is directed to the disk drive 21 in the second storage controller 20, the first storage controller 10 transmits the data frame to the second storage controller 20, and the second storage controller 20 executes a data input/output process directed to the disk drive 21 (S1031). In either cases, after the data input/output process has been completed, a report of completion is transmitted if necessary (S1032).

[0067] In (S1014), if the combination of Source_ID, Destination_ID of the data frame A does not exist in the priority management table, the first storage controller 10 checks, according to the procedure described in the above basic operation, whether the data frame A is directed to the disk drive 14 in the first storage controller 10 or to the second storage controller 20 (S1041). Now, if the data frame A is directed to the first storage controller 10, the data frame is transmitted to the storage device control unit 12 (S1042) and the storage device control unit 12 executes a data input/output process for the data frame A (S1043) and transmits to the host computer 5 a data frame in which a report of completion is described (S1044).

[0068] On the other hand, if the data frame A is directed to the disk drive 21 in the second storage controller 20, the first storage controller 10 transmits the data frame A to the second storage controller 20 (S1051). When the second storage controller 20 receives the data frame A (S1052), it executes a data input/output process directed to the disk drive 21 to the data frame A (S1053) and transmits a data frame, in which a report of completion is described, to the host computer 5 through the first storage controller 10 (S1054).

[0069] As understood from the above description, the priority control of data frames in the host computer 5 is dedicatedly executed in the first storage controller 10. That is, from another point of view, the first storage controller 10 executes replacing the role of the second storage controller 20 the priority control of the second storage controller 20. That is, it is enough that a priority control function is implemented on the first storage controller 10, and the second storage controller 20 can enjoy a function corresponding to the priority control function without having a mechanism for the function.

[0070] Furthermore, the load of operation can be reduced for the second storage controller 20 by the amount to be created by providing a mechanism for the priority control. Yet furthermore, it becomes possible to provide a function of priority control in the forms including the disk drive 21 in the second storage controller 20 even if a storage controller having a lower process capacity than that of the first storage controller 10 is used as the second storage controller 20 and, therefore, the introduction cost and the operation cost can be reduced.

[0071] Yet furthermore, old types of storage controllers having lower performance compared to the newly introduced storage controllers can be fully used as second storage controllers and, therefore, effective uses of resource and reduction of operation cost are progressed. In addition, it is enough that the priority management table is stored only in the first storage controller 10 and, therefore, reduction of the load of maintenance created by the unified management is progressed.

[0072] Though the priority control function for one LUN has been described as above, it can be arranged that a priority control is executed directing a plurality of LUNs and physical ports in a process described in one data frame.

[Data Copy Control]

[0073] In a storage system of the present invention, a copy management of data is conducted in a way in which a copy of the data stored in a logic volume (hereinafter referred to as "main logic volume") to which the first storage controller 10 executes data input/output processes is also stored in a logic volume (hereinafter referred to as "sub-logic volume") in the second storage controller 20. The relation between the LUN in the main logic volume and the LUN in the sub-logic volume is stored in the control memory 113 in the first storage controller 10 as a pair management table. An example of the pair management table is shown in Fig. 11. However, the relation may be automatically set by a function of the first storage controller 10 or may be set in a manual operation by an operator etc.

[0074] Now, a process of the storage system while a data frame in which a data writing request directed to the main logic volume is described is transmitted from the host computer 5 to the first storage controller 10 will be described referring to a flow chart shown in Fig. 12.

[0075] When the first storage controller 10 receives the data frame (S1262) transmitted from the host computer 5 (S1261), it checks whether an LUN described in the data frame as FCP_LUN (91) exists in the pair management table (S1263, S1264).

[0076] If the LUN described in the data frame as FCP_LUN (91) does not exist in the pair management table, the first storage controller 10 transmits the received data frame to the storage device control unit 12 (S1265). When the storage device control unit 12 receives the data frame (S1266), it executes to the proper logic volume a data writing process corresponding to the SCSI command described in the data frame (S1267) and, when necessary, transmits to the host computer 5 a data frame in which a report of completion is described (S1268). Therefore, data has been written in the main logic volume.

[0077] In (S1264), if the LUN exists in the pair management table, the first storage controller 10 creates a new data frame in which the Destination_ID (81) and the FCP_LUN (91) of the data frame is changed such that they designate the sub-logic volume being the destination of copying (S1269). The creation of the new data frame is executed in a way in which, after the received data frame has been stored in the control memory 113, the copy of the data frame is stored in the control memory and the Destination_ID (81) and the FCP_LUN (91) in the copy of the data frame are changed to create a new data frame.

[0078] Next, the first storage controller 10 transmits the created new data frame to the second storage controller 20 by controlling the switching device 111 to cause the switching device 111 to do so (S1270). When the second storage controller 20 receives the data frame (S1271), it executes a data writing process corresponding to the SCSI command described in the FCP_CDB of the data frame, to the sub-logic volume described in the FCP_LUN of the data frame (S1272). In this way, data being a copy of the data described in the main logic volume has been also written in the sub-logic volume. After the data writing process has been executed, the second storage controller 20 transmits to the host computer 5 a data frame in which a report of completion is described if necessary (S1273). In this way, data has been written in the main logic volume and, simultaneously, the same data has been written in the sub-logic volume.

[0079] On the other hand, in (S1264), if the LUN described in the data frame as the FCP_LUN (91) does not exist in the pair management table, the first storage con-

troller 10, following the procedure described in the above basic operation, checks whether the data frame is directed to the disk drive 14 in the first storage controller 10 or is directed to the second storage controller 20 (S1281). If the data frame is directed to the first storage controller 10, the data frame is transmitted to the storage device control unit 12 and the data writing process about the data frame is executed (1282). A data frame in which a report of completion is described is transmitted to the host computer 5 if necessary (S1284).

[0080] If the data frame is directed to the disk drive 21 in the second storage controller 20, the first storage controller 10 transmits the data frame to the second storage controller 20 (S1285). When the second storage controller 20 receives the data frame (S1286), it executes a data writing process directed to the disk drive 21 (S1287). Then, the second storage controller 20 transmits to the host computer 5 through the first storage controller 10 a data frame in which a report of completion of the data writing process is described if necessary (S1288).

[0081] In the above processes, when the first storage controller 10 receives a data frame in which a data writing request directed to the main logic volume is described, the first storage controller 10 creates a data frame for copying the request and transmits it to the second storage controller 20.

[0082] According to this method, a mechanism of data copy management, in which data stored in the disk drive 14 in the first storage controller 10 is also stored in the disk drive 21 in the second storage controller 20, can be easily realized only by providing a mechanism for creating a data frame as described above to the first storage controller 10.

[0083] There are two schemes for data copy management known such as so-called "synchronized scheme" in which the first storage controller 10 having received the data frame notifies the host computer 5 of a report of completion after confirming the completion of the writing processes for both of the main logic volume and the sub-logic volume, and so-called "non-synchronized scheme" in which the first storage controller 10 notifies the host computer 5 of the report of completion only when the writing process to the main logic volume has been completed regardless of whether the writing process to the sub-logic volume has been completed or not. However, the mechanism of the above-described embodiment can be applied to the case operated in either the "synchronized scheme" or the "non-synchronized scheme".

[0084] Although the storage system control method, etc., according to the present invention have been set forth by way of several embodiments, the above embodiments are only to facilitate the understanding of the present invention but are not intended to limit the present invention. It will be appreciated that the present invention can be variously changed or improved and that the present invention covers equivalents thereof. Only one of the functions of access limit, priority, data copy control described above may be applied or any combination of two or more of these may be applied.

[0085] The communication protocol used for the first communication means or the second communication means is not limited to the Fiber Channel Protocol, but instead other protocols may be used.

[0086] In the above embodiments, a disk array apparatus has been described as an example of the storage apparatus. The present invention is however applicable to other storage apparatuses than the disk array apparatus, such as a semiconductor disk apparatus.

[0087] Although the present invention has been described above based on the embodiments, it is to be noted that the present invention shall not be limited to the embodiments stated herein, and that various modifications can be made.

Claims

1. A method of controlling the data access in a virtualization system coupled to a plurality of host systems (5) and at least one storage system (10, 20), wherein at least one of the host systems (5) sends a write request to a volume in the virtualization system, the storage system (10, 20) having a plurality of disk drives (14, 21) and a plurality of logical volumes related to the disk drives (14, 21), the method comprising:

storing a first relationship information between a first logical volume and a second volume;
storing a second relationship information between a specified one of the host systems (5) and the second volume;
storing a third relationship information for identifying a pair relationship between the second volume and a first volume;
receiving a first write request from a one of the host systems (5) targeted to the first volume;
transferring data of the first write request to the first volume;
copying data of the first write request to the second volume based on said third relationship information, the second volume being a virtual volume for transferring the copy of data of the first write request to the first logical volume of the storage system (10, 20);
transferring the copy of data of the first write request to the storage system (10, 20) based on said first relationship information;
receiving a second write request from a second one of the host systems (5) targeted to the second volume;
restricting the data transfer of the second write request to the second volume based on said second relationship information, which includes a relationship between the second host system (5) and the second volume, if the second host

system (5) is not allowed to access the second volume; and
transferring the data of the second write request to the second volume if the second host system (5) is allowed to access the second volume.

2. A method of controlling the data access in a virtualization system coupled to a plurality of host systems (5) and at least one storage system (10, 20), wherein at least one of the host systems (5) sends a write request to a volume in the virtualization system, the storage system (10, 20) having a plurality of disk drives (14, 21) and a plurality of logical volumes related to the disk drives (14, 21), the method comprising:

storing a first relationship information between a first logical volume and a second volume;
storing a second relationship information between a specified one of the host systems (5) and the second volume,
storing a third relationship information for identifying a pair relationship between the second volume and a first volume,
receiving a first write request from a one of the host systems (5) targeted to the first volume;
transferring data of the first write request to the first volume;
copying data of the first write request to the second volume based on said third relationship information, the second volume being a virtual volume for transferring the copy of data of the first write request to the first logical volume of the storage system (10, 20);
transferring the copy of data of the first write request to the storage system (10, 20) based on said first relationship information;
receiving a second write request from a second one the host systems (5) targeted to the second volume;
determining whether or not the second host system (5) is allowed to access the second volume, based on said second relationship information, which includes a relationship between the second host system (5) and the second volume;
restricting the data transfer of the second write request to the second volume if the second host system (5) is not allowed to access the second volume; and
transferring the data of the second write request to the second volume if the second host system (5) is allowed to access the second volume.

3. The method of claim 1 or 2, wherein:

the virtualization system includes a data switching part (11) and a control part (12); and
the data switching part (11) performs the steps of receiving the first write request, transferring data of the first write request to the first volume, transferring the copy of data of the first write request to the storage system (10, 20), receiving the second write request, restricting the data transfer of the second write request to the second volume, and transferring the data of the second write request to the second volume.

4. The method of claim 1 or 2, wherein each of the first volume, the second volume and the first logical volume is identified by a Logical Unit Number (LUN).

5. The method of claim 1 or 2, wherein the first relationship information relates the second volume to both the first logical volume and a second logical volume of the storage system (10, 20).

6. A virtualization system coupled to a plurality of host systems (5) and at least one storage system (10, 20), wherein at least one of the host systems (5) sends a write request to a volume in the virtualization system, the storage system (10, 20) having a plurality of disk drives (14, 21) and a plurality of logical volumes related to the disk drives (14, 21), the virtualization system comprising:

a plurality ports coupled to the host systems (5) and the storage system (10, 20);
a data switching part (11); and
a control part (12);
wherein the data switching part (11) and/or the control part (12) is/are adapted to:

store a first relationship information between a first logical volume and a second volume;
store a second relationship information between a specified one of the host systems (5) and the second volume;
store a third relationship information for identifying a pair relationship between the second volume and a first volume;
receive a first write request from a first one of the host systems (5) targeted to the first volume;
transfer data of the first write request to the first volume;
copy data of the first write request to the second volume based on said third relationship information, the second volume being a virtual volume for transferring the copy of data of the first write request to the first logical volume of the storage system (10, 20);
transfer the copy of data of the first write request to the storage system (10, 20), based on said first relationship information;
receive a second write request from a sec-

ond one of the host systems (5) targeted to the second volume;
restrict the data transfer of the second write request to the second volume based on said second relationship information, which includes a relationship between the second host system (5) and the second volume, if the second host system (5) is not allowed to access to the second volume; and
transfer the data of the second write request to the second volume if the second host system (5) is allowed to access to the second volume.

7. A virtualization system coupled to a plurality of host systems (5) and at least one storage system (10, 20), wherein at least one of the host systems (5) sends a write request to a volume in the virtualization system, the storage system (10, 20) having a plurality of disk drives (14, 21) and a plurality of logical volumes related to the disk drives (14, 21), the virtualization system comprising:

a plurality ports coupled to the host systems (5) and the storage system (10, 20);
a data switching part (11); and
a control part (12);
wherein the data switching part (11) and/or the control part (12) is/are adapted to:

store a first relationship information between a first logical volume and a second volume;
store a second relationship information between a specified one of the host systems (5) and the second volume;
store a third relationship information for identifying a pair relationship between the second volume and a first volume;
receive a first write request from a first one of the host systems (5) targeted to the first volume;
transfer data of the first write request to the first volume;
copy data of the first write request to the second volume based on said third relationship information, the second volume being a virtual volume for transferring the copy of data of the first write request to the first logical volume of the storage system (10, 20);
transfer the copy of data of the first write request to the storage system (10, 20), based on said first relationship information;
receive a second write request from a second one of the host systems (5) targeted to the second volume;
determine whether or not the second host system (5) is allowed to access the second volume, based on said second relationship information, which includes a relationship between the second host system (5) and the second volume;
restrict the data transfer of the second write request to the second volume if the second host system (5) is not allowed to access to the second volume; and
transfer the data of the second write request to the second volume if the second host system (5) is allowed to access to the second volume.

8. The system of claim 6 or 7, wherein the data switching part (11) is adapted to receive the first write request, transfer data of the first write request to the first volume, transfer the copy of data of the first write request to the storage system (10, 20), receive the second write request, restrict the data transfer of the second write request to the second volume, and transfer the data of the second write request to the second volume.

9. The system of claim 6 or 7, wherein each of the first volume, the second volume and the first logical volume is identified by a Logical Unit Number (LUN).

10. The system of claim 6 or 7, wherein the first relationship information relates the second volume to both the first logical volume and a second logical volume of the storage system (10, 20).

Patentansprüche

1. Verfahren zur Datenzugriffssteuerung in einem mit mehreren Hostsystemen (5) und mindestens einem Speichersystem (10, 20) gekoppelten Virtualisierungssystem, wobei mindestens eines der Hostsysteme (5) an ein Volumen in dem Virtualisierungssystem einen Schreibbefehl sendet, und wobei das Speichersystem (10, 20) mehrere Plattenlaufwerke (14, 21) und mehrere auf diese bezogene logische Volumina aufweist, wobei in dem Verfahren:

eine erste Beziehungsinformation zwischen einem ersten logischen Volumen und einem zweiten Volumen gespeichert wird,
eine zweite Beziehungsinformation zwischen einem bestimmten der Hostsysteme (5) und dem zweiten Volumen gespeichert wird,
eine dritte Beziehungsinformation zum Identifizieren einer Paarbeziehung zwischen dem zweiten Volumen und einem ersten Volumen gespeichert wird,
ein von einem der Hostsysteme (5) ausgehender, auf das erste Volumen gerichteter erster Schreibbefehl empfangen wird,

Daten des ersten Schreibbefehls in das erste Volumen übertragen werden,
aufgrund der dritten Beziehungsinformation Daten des ersten Schreibbefehls in das zweite Volumen kopiert werden, wobei das zweite Volumen ein virtuelles Volumen zum Übertragen der Datenkopie des ersten Schreibbefehls in das erste logische Volumen des Speichersystems (10, 20) ist,
aufgrund der ersten Beziehungsinformation die Datenkopie des ersten Schreibbefehls in das Speichersystem (10, 20) übertragen wird,
ein von einem zweiten der Hostsysteme (5) ausgehender, auf das zweite Volumen gerichteter zweiter Schreibbefehl empfangen wird,
aufgrund der eine Beziehung zwischen dem zweiten Hostsystem (5) und dem zweiten Volumen enthaltenden zweiten Beziehungsinformation die Datenübertragung des zweiten Schreibbefehls auf das zweite Volumen beschränkt wird, falls das zweite Hostsystem (5) keine Berechtigung zum Zugriff auf das zweite Volumen hat, und
die Daten des zweiten Schreibbefehls auf das zweite Volumen übertragen werden, falls das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat.

2. Verfahren zur Datenzugriffssteuerung in einem mit mehreren Hostsystemen (5) und mindestens einem Speichersystem (10, 20) gekoppelten Virtualisierungssystem, wobei mindestens eines der Hostsysteme (5) an ein Volumen in dem Virtualisierungssystem einen Schreibbefehl sendet, und wobei das Speichersystem (10, 20) mehrere Plattenlaufwerke (14, 21) und mehrere auf diese bezogene logische Volumina aufweist, wobei in dem Verfahren:

eine erste Beziehungsinformation zwischen einem ersten logischen Volumen und einem zweiten Volumen gespeichert wird,
eine zweite Beziehungsinformation zwischen einem bestimmten der Hostsysteme (5) und dem zweiten Volumen gespeichert wird,
eine dritte Beziehungsinformation zum Identifizieren einer Paarbeziehung zwischen dem zweiten Volumen und einem ersten Volumen gespeichert wird,
ein von einem der Hostsysteme (5) ausgehender, auf das erste Volumen gerichteter erster Schreibbefehl empfangen wird,
Daten des ersten Schreibbefehls in das erste Volumen übertragen werden,
aufgrund der dritten Beziehungsinformation Daten des ersten Schreibbefehls in das zweite Volumen kopiert werden, wobei das zweite Volumen ein virtuelles Volumen zum Übertragen der Datenkopie des ersten Schreibbefehls in das erste logische Volumen des Speichersystems (10, 20) ist,
aufgrund der ersten Beziehungsinformation die Datenkopie des ersten Schreibbefehls in das Speichersystem (10, 20) übertragen wird,
ein von einem zweiten der Hostsysteme (5) ausgehender, auf das zweite Volumen gerichteter zweiter Schreibbefehl empfangen wird,
aufgrund der eine Beziehung zwischen dem zweiten Hostsystem (5) und dem zweiten Volumen enthaltenden zweiten Beziehungsinformation bestimmt wird, ob das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat oder nicht,
die Datenübertragung des zweiten Schreibbefehls auf das zweite Volumen beschränkt wird, falls das zweite Hostsystem (5) keine Berechtigung zum Zugriff auf das zweite Volumen hat, und
die Daten des zweiten Schreibbefehls auf das zweite Volumen übertragen werden, falls das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat.

3. Verfahren nach Anspruch 1 oder 2, wobei das Virtualisierungssystem einen Datenvermittlungsteil (11) und einen Steuerteil (12) aufweist und der Datenvermittlungsteil (11) den Empfang des ersten Schreibbefehls, die Übertragung von Daten des ersten Schreibbefehls an das erste Volumen, die Übertragung der Datenkopie des ersten Schreibbefehls an das Speichersystem (10, 20), den Empfang des zweiten Schreibbefehls, die Beschränkung der Datenübertragung des zweiten Schreibbefehls auf das erste Volumen sowie die Übertragung der Daten des zweiten Schreibbefehls an das zweite Volumen durchführt.

4. Verfahren nach Anspruch 1 oder 2, wobei das erste Volumen, das zweite Volumen und das dritte Volumen jeweils durch eine Logikeinheitennummer (LUN) identifiziert sind.

5. Verfahren nach Anspruch 1 oder 2, wobei die erste Beziehungsinformation das zweite Volumen auf sowohl das erste logische Volumen als auch ein zweites logisches Volumen des Speichersystems (10, 20) bezieht.

6. Virtualisierungssystem, das mit mehreren Hostsystemen (5) und mindestens einem Speichersystem (10, 20) gekoppelt ist, wobei mindestens eines der Hostsysteme (5) an ein Volumen in dem Virtualisierungssystem einen Schreibbefehl sendet, wobei das Speichersystem (10, 20) mehrere Plattenlaufwerke (14, 21) und mehrere auf diese bezogene logische Volumina aufweist, und wobei das Virtualisierungssystem aufweist:

mehrere mit den Hostsystemen (5) und dem Speichersystem (10, 20) gekoppelte Ports,
einen Datenvermittlungsteil (11) und
einen Steuerteil (12),
wobei der Datenvermittlungsteil (11) und/oder der Steuerteil (12) dazu ausgelegt ist/sind:

eine erste Beziehungsinformation zwischen einem ersten logischen Volumen und einem zweiten Volumen zu speichern,
eine zweite Beziehungsinformation zwischen einem bestimmten der Hostsysteme (5) und dem zweiten Volumen zu speichern,
eine dritte Beziehungsinformation zum Identifizieren einer Paarbeziehung zwischen dem zweiten Volumen und einem ersten Volumen zu speichern,
einen von einem der Hostsysteme (5) ausgehenden, auf das erste Volumen gerichteten ersten Schreibbefehl zu empfangen,
Daten des ersten Schreibbefehls in das erste Volumen zu übertragen,
aufgrund der dritten Beziehungsinformation Daten des ersten Schreibbefehls in das zweite Volumen zu kopieren, wobei das zweite Volumen ein virtuelles Volumen zum Übertragen der Datenkopie des ersten Schreibbefehls in das erste logische Volumen des Speichersystems (10, 20) ist,
aufgrund der ersten Beziehungsinformation die Datenkopie des ersten Schreibbefehls in das Speichersystem (10, 20) zu übertragen,
einen von einem zweiten der Hostsysteme (5) ausgehenden, auf das zweite Volumen gerichteten zweiten Schreibbefehl zu empfangen,
aufgrund der eine Beziehung zwischen dem zweiten Hostsystem (5) und dem zweiten Volumen enthaltenden zweiten Beziehungsinformation die Datenübertragung des zweiten Schreibbefehls auf das zweite Volumen zu beschränken, falls das zweite Hostsystem (5) keine Berechtigung zum Zugriff auf das zweite Volumen hat, und
die Daten des zweiten Schreibbefehls auf das zweite Volumen zu übertragen, falls das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat.

7. Virtualisierungssystem, das mit mehreren Hostsystemen (5) und mindestens einem Speichersystem (10, 20) gekoppelt ist, wobei mindestens eines der Hostsysteme (5) an ein Volumen in dem Virtualisierungssystem einen Schreibbefehl sendet, wobei das Speichersystem (10, 20) mehrere Plattenlaufwerke (14, 21) und mehrere auf diese bezogene logische Volumina aufweist, und wobei das Virtualisierungssystem aufweist:

mehrere mit den Hostsystemen (5) und dem Speichersystem (10, 20) gekoppelte Ports,
einen Datenvermittlungsteil (11) und
einen Steuerteil (12),
wobei der Datenvermittlungsteil (11) und/oder der Steuerteil (12) dazu ausgelegt ist/sind:

eine erste Beziehungsinformation zwischen einem ersten logischen Volumen und einem zweiten Volumen zu speichern,
eine zweite Beziehungsinformation zwischen einem bestimmten der Hostsysteme (5) und dem zweiten Volumen zu speichern,
eine dritte Beziehungsinformation zum Identifizieren einer Paarbeziehung zwischen dem zweiten Volumen und einem ersten Volumen zu speichern,
einen von einem der Hostsysteme (5) ausgehenden, auf das erste Volumen gerichteten ersten Schreibbefehl zu empfangen,
Daten des ersten Schreibbefehls in das erste Volumen zu übertragen,
aufgrund der dritten Beziehungsinformation Daten des ersten Schreibbefehls in das zweite Volumen zu kopieren, wobei das zweite Volumen ein virtuelles Volumen zum Übertragen der Datenkopie des ersten Schreibbefehls in das erste logische Volumen des Speichersystems (10, 20) ist,
aufgrund der ersten Beziehungsinformation die Datenkopie des ersten Schreibbefehls in das Speichersystem (10, 20) zu übertragen,
einen von einem zweiten der Hostsysteme (5) ausgehenden, auf das zweite Volumen gerichteten zweiten Schreibbefehl zu empfangen,
aufgrund der eine Beziehung zwischen dem zweiten Hostsystem (5) und dem zweiten Volumen enthaltenden zweiten Beziehungsinformation zu bestimmen, ob das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat oder nicht,
die Datenübertragung des zweiten Schreibbefehls auf das zweite Volumen zu beschränken, falls das zweite Hostsystem (5) keine Berechtigung zum Zugriff auf das zweite Volumen hat, und
die Daten des zweiten Schreibbefehls auf das zweite Volumen zu übertragen, falls das zweite Hostsystem (5) Berechtigung zum Zugriff auf das zweite Volumen hat.

8. System nach Anspruch 6 oder 7, wobei der Datenvermittlungsteil (11) dazu ausgelegt ist, den ersten

Schreibbefehl zu empfangen, Daten des ersten Schreibbefehls an das erste Volumen zu übertragen, die Datenkopie des ersten Schreibbefehls an das Speichersystem (10, 20) zu übertragen, den zweiten Schreibbefehl zu empfangen, die Datenübertragung des zweiten Schreibbefehls auf das erste Volumen zu beschränken sowie die Daten des zweiten Schreibbefehls an das zweite Volumen zu übertragen.

9. System nach Anspruch 6 oder 7, wobei das erste Volumen, das zweite Volumen und das dritte Volumen jeweils durch eine Logikeinheitennummer (LUN) identifiziert sind.

10. System nach Anspruch 6 oder 7, wobei die erste Beziehungsinformation das zweite Volumen auf sowohl das erste logische Volumen als auch ein zweites logisches Volumen des Speichersystems (10, 20) bezieht.

Revendications

1. Procédé pour commander l'accès à des données dans un système de virtualisation couplé à une pluralité de systèmes hôtes (5) et à au moins un système de mémorisation (10, 20), dans lequel au moins l'un des systèmes hôtes (5) envoie une demande d'écriture à un volume dans le système de virtualisation, le système de mémorisation (10, 20) ayant une pluralité d'unités de disque (14, 21) et une pluralité de volumes logiques associés aux unités de disque (14, 21), le procédé comportant les étapes consistant à :

mémoriser des premières informations de relation entre un premier volume logique et un second volume,
mémoriser des deuxièmes informations de relation entre un système spécifié parmi les systèmes hôtes (5) et le second volume,
mémoriser des troisièmes informations de relation pour identifier une relation de paire entre le second volume et un premier volume,
recevoir une première demande d'écriture en provenance de l'un des systèmes hôtes (5) destinée à un premier volume,
transférer des données de la première demande d'écriture vers le premier volume,
copier des données de la première demande d'écriture dans le second volume sur la base desdites troisièmes informations de relation, le second volume étant un volume virtuel pour transférer la copie de données de la première demande d'écriture vers le premier volume logique du système de mémorisation (10, 20),
transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20) sur la base desdites premières informations de relation,
recevoir une seconde demande d'écriture en provenance d'un second système parmi les systèmes hôtes (5) destinée au second volume,
limiter le transfert de données de la seconde demande d'écriture vers le second volume sur la base desdites deuxièmes informations de relation, lesquelles incluent une relation entre le second système hôte (5) et le second volume, si le second système hôte (5) n'est pas autorisé à accéder au second volume, et
transférer les données de la seconde demande d'écriture vers le second volume si le second système hôte (5) est autorisé à accéder au second volume.

2. Procédé pour commander l'accès à des données dans un système de virtualisation couplé à une pluralité de systèmes hôtes (5) et à au moins un système de mémorisation (10, 20), dans lequel au moins l'un des systèmes hôtes (5) envoie une demande d'écriture à un volume dans le système de virtualisation, le système de mémorisation (10, 20) ayant une pluralité d'unités de disque (14, 21) et une pluralité de volumes logiques associés aux unités de disque (14, 21), le procédé comportant les étapes consistant à :

mémoriser des premières informations de relation entre un premier volume logique et un second volume,
mémoriser des deuxièmes informations de relation entre un système spécifié parmi les systèmes hôtes (5) et le second volume,
mémoriser des troisièmes informations de relation pour identifier une relation de paire entre le second volume et un premier volume,
recevoir une première demande d'écriture en provenance de l'un des systèmes hôtes (5) destinée à un premier volume,
transférer des données de la première demande d'écriture vers le premier volume,
copier des données de la première demande d'écriture dans le second volume sur la base desdites troisièmes informations de relation, le second volume étant un volume virtuel pour transférer la copie de données de la première demande d'écriture vers le premier volume logique du système de mémorisation (10, 20),
transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20) sur la base desdites premières informations de relation,
recevoir une seconde demande d'écriture en provenance d'un second système parmi les systèmes hôtes (5) destinée au second volume,

déterminer si oui ou non le second système hôte (5) est autorisé à accéder au second volume, sur la base desdites deuxièmes informations de relation, lesquelles incluent une relation entre le second système hôte (5) et le second volume,
limiter le transfert de données de la seconde demande d'écriture vers le second volume si le second système hôte (5) n'est pas autorisé à accéder au second volume, et
transférer les données de la seconde demande d'écriture vers le second volume si le second système hôte (5) est autorisé à accéder au second volume.

3. Procédé selon la revendication 1 ou 2, dans lequel :

le système de virtualisation inclut une partie de commutation de données (11) et une partie de commande (12), et
la partie de commutation de données (11) exécute les étapes consistant à recevoir la première demande d'écriture, transférer des données de la première demande d'écriture vers le premier volume, transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20), recevoir la seconde demande d'écriture, limiter le transfert de données de la seconde demande d'écriture vers le second volume, et transférer les données de la seconde demande d'écriture vers le second volume.

4. Procédé selon la revendication 1 ou 2, dans lequel chaque volume parmi le premier volume, le second volume et le premier volume logique est identifié par un numéro d'unité logique (LUN).

5. Procédé selon la revendication 1 ou 2, dans lequel les premières informations de relation concernent le second volume à la fois du premier volume logique et d'un second volume logique du système de mémorisation (10, 20).

6. Système de virtualisation couplé à une pluralité de systèmes hôtes (5) et à au moins un système de mémorisation (10, 20), dans lequel au moins l'un des systèmes hôtes (5) envoie une demande d'écriture à un volume dans le système de virtualisation, le système de mémorisation (10, 20) ayant une pluralité d'unités de disque (14, 21) et une pluralité de volumes logiques associés aux unités de disque (14, 21), le système de virtualisation, comportant :

une pluralité de ports couplés aux systèmes hôtes (5) et au système de mémorisation (10, 20),
une partie de commutation de données (11), et
une partie de commande (12),
dans lequel la partie de commutation de données (11) et/ou la partie de commande (12) est/sont adaptée(s) pour :

mémoriser des premières informations de relation entre un premier volume logique et un second volume,
mémoriser des deuxièmes informations de relation entre un système spécifié parmi les systèmes hôtes (5) et le second volume,
mémoriser des troisièmes informations de relation pour identifier une relation de paire entre le second volume et un premier volume,
recevoir une première demande d'écriture en provenance de l'un des systèmes hôtes (5) destinée à un premier volume,
transférer des données de la première demande d'écriture vers le premier volume,
copier des données de la première demande d'écriture dans le second volume sur la base desdites troisièmes informations de relation, le second volume étant un volume virtuel pour transférer la copie de données de la première demande d'écriture vers le premier volume logique du système de mémorisation (10, 20),
transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20), sur la base desdites premières informations de relation,
recevoir une seconde demande d'écriture en provenance d'un second système parmi les systèmes hôtes (5) destinée au second volume,
limiter le transfert de données de la seconde demande d'écriture vers le second volume sur la base desdites deuxièmes informations de relation, lesquelles incluent une relation entre le second système hôte (5) et le second volume, si le second système hôte (5) n'est pas autorisé à accéder au second volume, et
transférer les données de la seconde demande d'écriture vers le second volume si le second système hôte (5) est autorisé à accéder au second volume.

7. Système de virtualisation couplé à une pluralité de systèmes hôtes (5) et à au moins un système de mémorisation (10, 20),dans lequel au moins l'un des systèmes hôtes (5) envoie une demande d'écriture à un volume dans le système de virtualisation, le système de mémorisation (10, 20) ayant une pluralité d'unités de disque (14, 21) et une pluralité de volumes logiques associés aux unités de disque (14, 21), le système de virtualisation, comportant :

une pluralité de ports couplés aux systèmes hô-

tes (5) et au système de mémorisation (10, 20), une partie de commutation de données (11), et une partie de commande (12),
dans lequel la partie de commutation de données (11) et/ou la partie de commande (12) est/sont adaptée(s) pour :

mémoriser des premières informations de relation entre un premier volume logique et un second volume,
mémoriser des deuxièmes informations de relation entre un système spécifié parmi les systèmes hôtes (5) et le second volume,
mémoriser des troisièmes informations de relation pour identifier une relation de paire entre le second volume et un premier volume,
recevoir une première demande d'écriture en provenance de l'un des systèmes hôtes (5) destinée à un premier volume,
transférer des données de la première demande d'écriture vers le premier volume,
copier des données de la première demande d'écriture dans le second volume sur la base desdites troisièmes informations de relation, le second volume étant un volume virtuel pour transférer la copie de données de la première demande d'écriture vers le premier volume logique du système de mémorisation (10, 20),
transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20), sur la base desdites premières informations de relation,
recevoir une seconde demande d'écriture en provenance d'un second système parmi les systèmes hôtes (5) destinée au second volume,
déterminer si oui ou non le second système hôte (5) est autorisé à accéder au second volume, sur la base desdites deuxièmes informations de relation, lesquelles incluent une relation entre le second système hôte (5) et le second volume,
limiter le transfert de données de la seconde demande d'écriture vers le second volume si le second système hôte (5) n'est pas autorisé à accéder au second volume, et
transférer les données de la seconde demande d'écriture vers le second volume si le second système hôte (5) est autorisé à accéder au second volume.

8. Système selon la revendication 6 ou 7, dans lequel la partie de commutation de données (11) est adaptée pour recevoir la première demande d'écriture, transférer des données de la première demande d'écriture vers le premier volume, transférer la copie de données de la première demande d'écriture vers le système de mémorisation (10, 20), recevoir la seconde demande d'écriture, limiter le transfert de données de la seconde demande d'écriture vers le second volume, et transférer les données de la seconde demande d'écriture vers le second volume.

9. Système selon la revendication 6 ou 7, dans lequel chaque volume parmi le premier volume, le second volume et le premier volume logique est identifié par un numéro d'unité logique (LUN).

10. Système selon la revendication 6 ou 7, dans lequel les premières informations de relation concernent le second volume à la fois du premier volume logique et d'un second volume logique du système de mémorisation (10, 20).

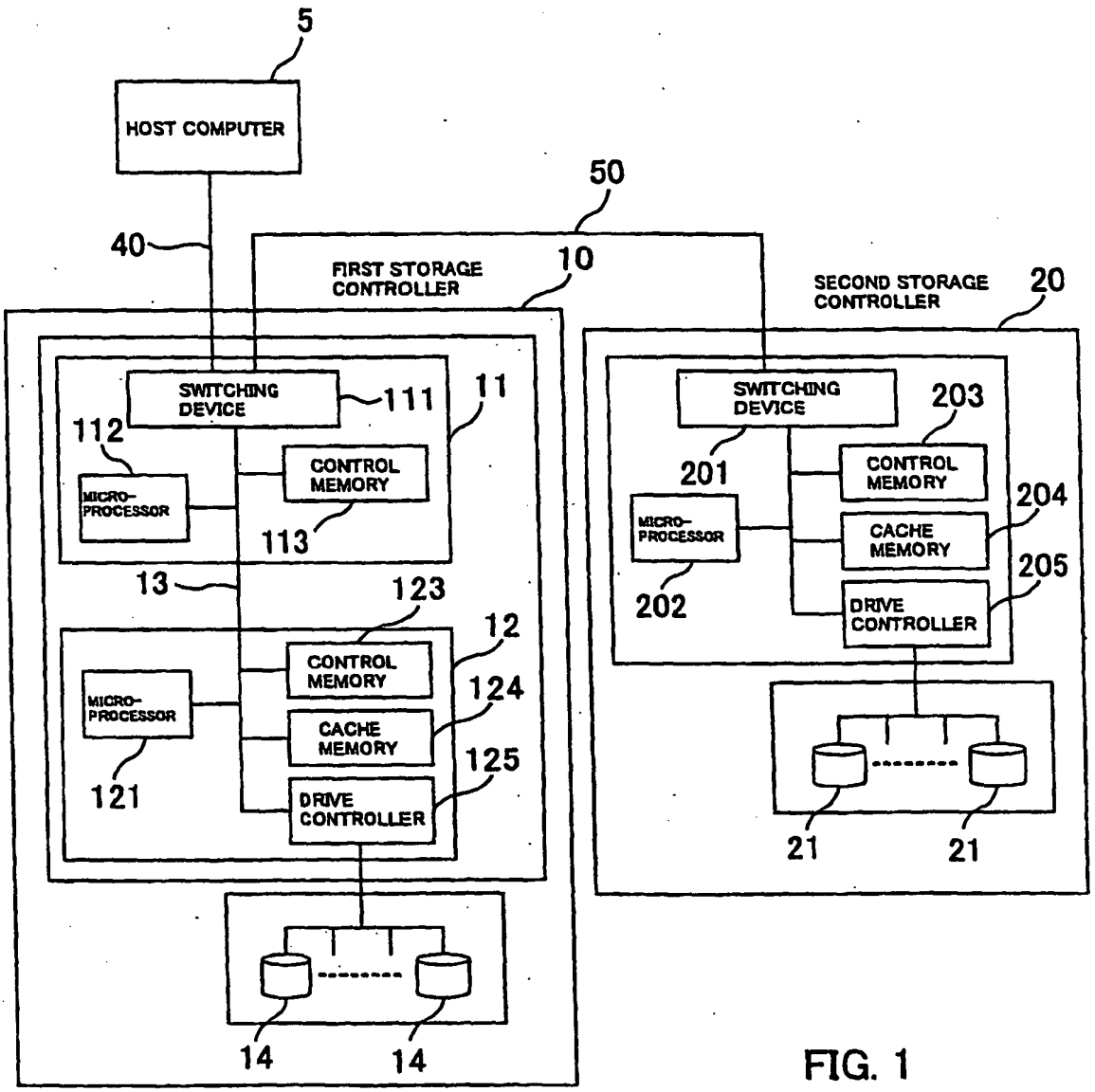

5
HOST COMPUTER
50
40
FIRST STORAGE CONTROLLER
10
SECOND STORAGE CONTROLLER
20
SWITCHING DEVICE
111
11
112
CONTROL MEMORY
MICRO-PROCESSOR
113
13
123
12
CONTROL MEMORY
124
MICRO-PROCESSOR
CACHE MEMORY
125
121
DRIVE CONTROLLER
14
14
SWITCHING DEVICE
203
201
CONTROL MEMORY
204
MICRO-PROCESSOR
CACHE MEMORY
205
202
DRIVE CONTROLLER
21
21

FIG. 1

FIG. 2

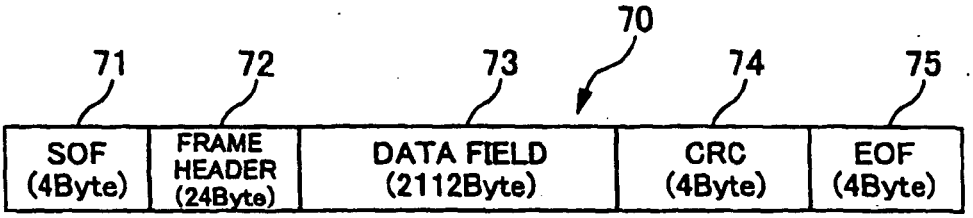


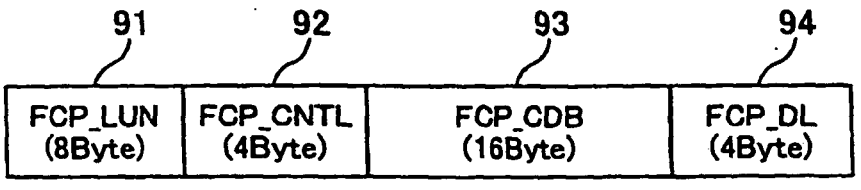
FIG. 3

Byte \ Bit	31–24	23–16	15–8	7–0
0	R_CTL	DESTINATION_ID		
1	Reserved	SOURCE_ID		
2	TYPE	F_CTL		
3	SEQ_ID	DF_CTL	SEQ_CNT	
4	OX_ID		RX_ID	
5	PARAMETER			

81

82

FIG. 4



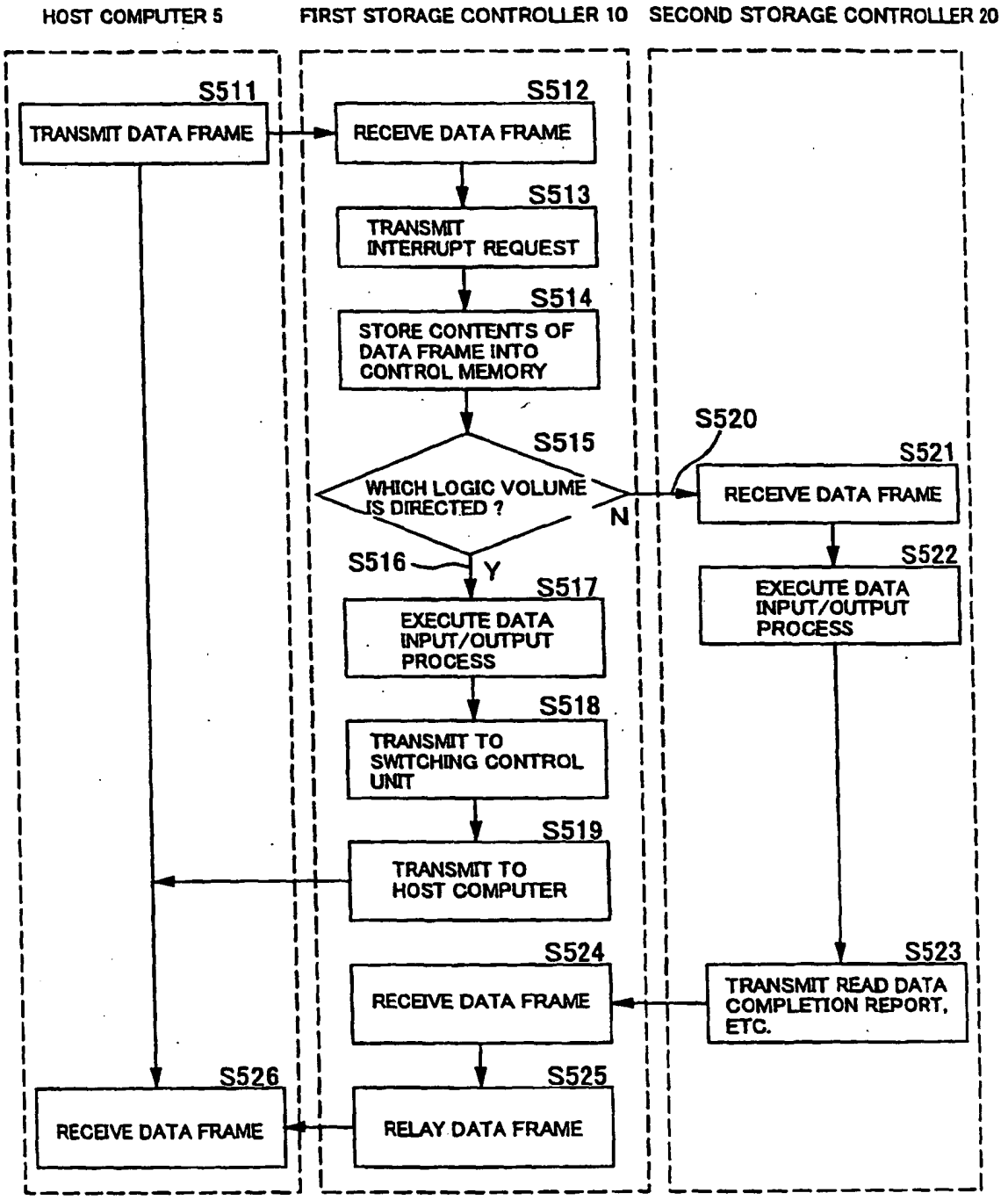

HOST COMPUTER 5
FIRST STORAGE CONTROLLER 10
SECOND STORAGE CONTROLLER 20
S511
TRANSMIT DATA FRAME
S512
RECEIVE DATA FRAME
S513
TRANSMIT INTERRUPT REQUEST
S514
STORE CONTENTS OF DATA FRAME INTO CONTROL MEMORY
S515
WHICH LOGIC VOLUME IS DIRECTED ?
N
S520
S521
RECEIVE DATA FRAME
S516
Y
S517
EXECUTE DATA INPUT/OUTPUT PROCESS
S522
EXECUTE DATA INPUT/OUTPUT PROCESS
S518
TRANSMIT TO SWITCHING CONTROL UNIT
S519
TRANSMIT TO HOST COMPUTER
S524
RECEIVE DATA FRAME
S523
TRANSMIT READ DATA COMPLETION REPORT, ETC.
S526
RECEIVE DATA FRAME
S525
RELAY DATA FRAME

FIG. 5

FIG. 6

LUN CONTROL TABLE

LUN	FIRST STORAGE CONTROLLER	SECOND STORAGE CONTROLLER
0001	O	
0002	O	
0003		O
0004		O

FIG. 7

ACCESS LIMIT MANAGEMENT TABLE

Destination_ID	Source_ID	LUN
522	124	0001
522	124	0002
525	153	0014
525	153	0015
•	•	•
•	•	•

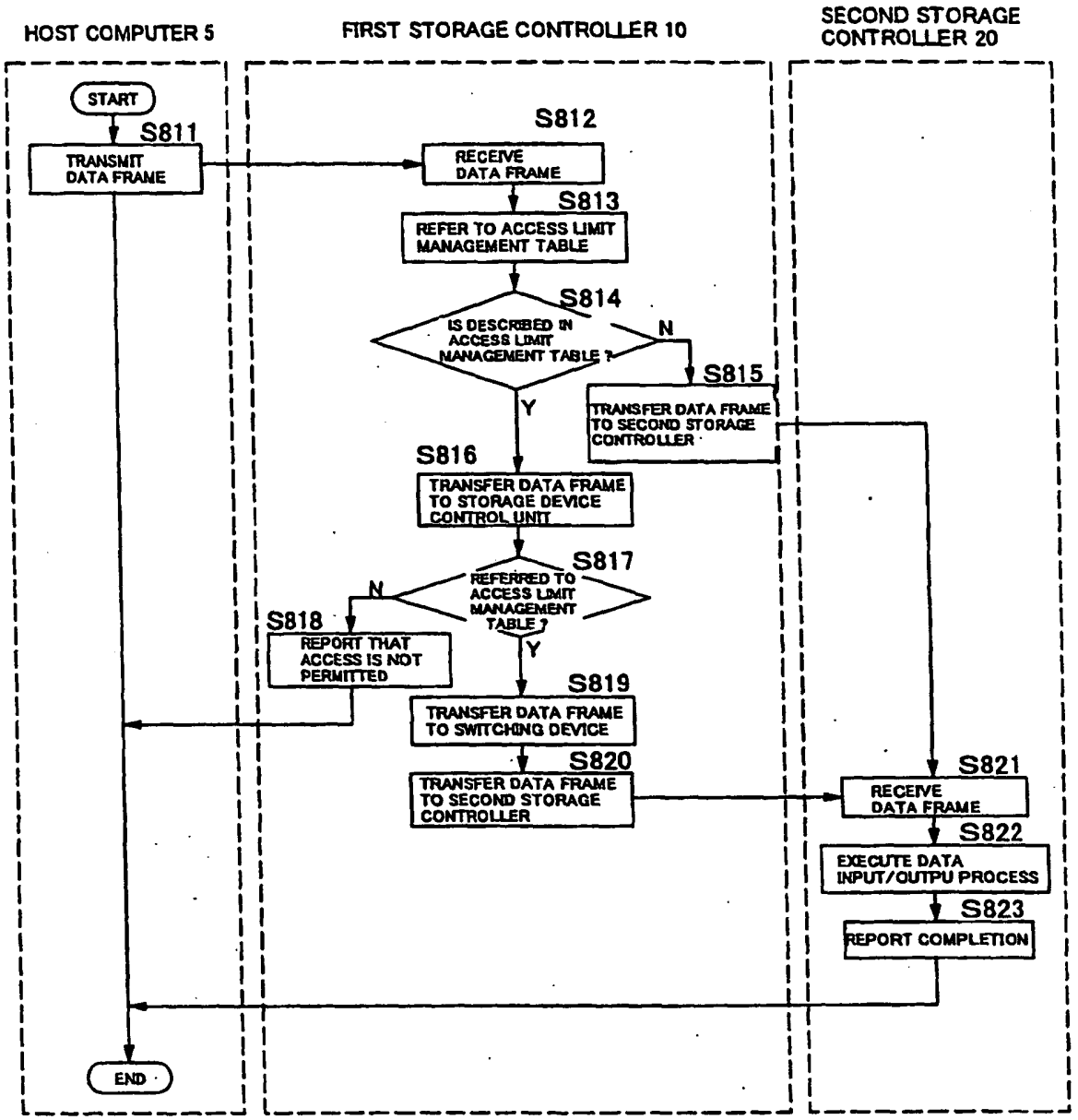

HOST COMPUTER 5
FIRST STORAGE CONTROLLER 10
SECOND STORAGE CONTROLLER 20
START
S811
TRANSMIT DATA FRAME
S812
RECEIVE DATA FRAME
S813
REFER TO ACCESS LIMIT MANAGEMENT TABLE
S814
IS DESCRIBED IN ACCESS LIMIT MANAGEMENT TABLE ?
N
Y
S815
TRANSFER DATA FRAME TO SECOND STORAGE CONTROLLER
S816
TRANSFER DATA FRAME TO STORAGE DEVICE CONTROL UNIT
S817
REFERRED TO ACCESS LIMIT MANAGEMENT TABLE ?
N
Y
S818
REPORT THAT ACCESS IS NOT PERMITTED
S819
TRANSFER DATA FRAME TO SWITCHING DEVICE
S820
TRANSFER DATA FRAME TO SECOND STORAGE CONTROLLER
S821
RECEIVE DATA FRAME
S822
EXECUTE DATA INPUT/OUTPU PROCESS
S823
REPORT COMPLETION
END

FIG. 8

FIG. 9

PRIORITY MANAGEMENT TABLE

Destination_ID	Source_ID	DELAY TIME(SEC.)
522	124	100
523	124	150
525	153	100
•	•	•
•	•	•

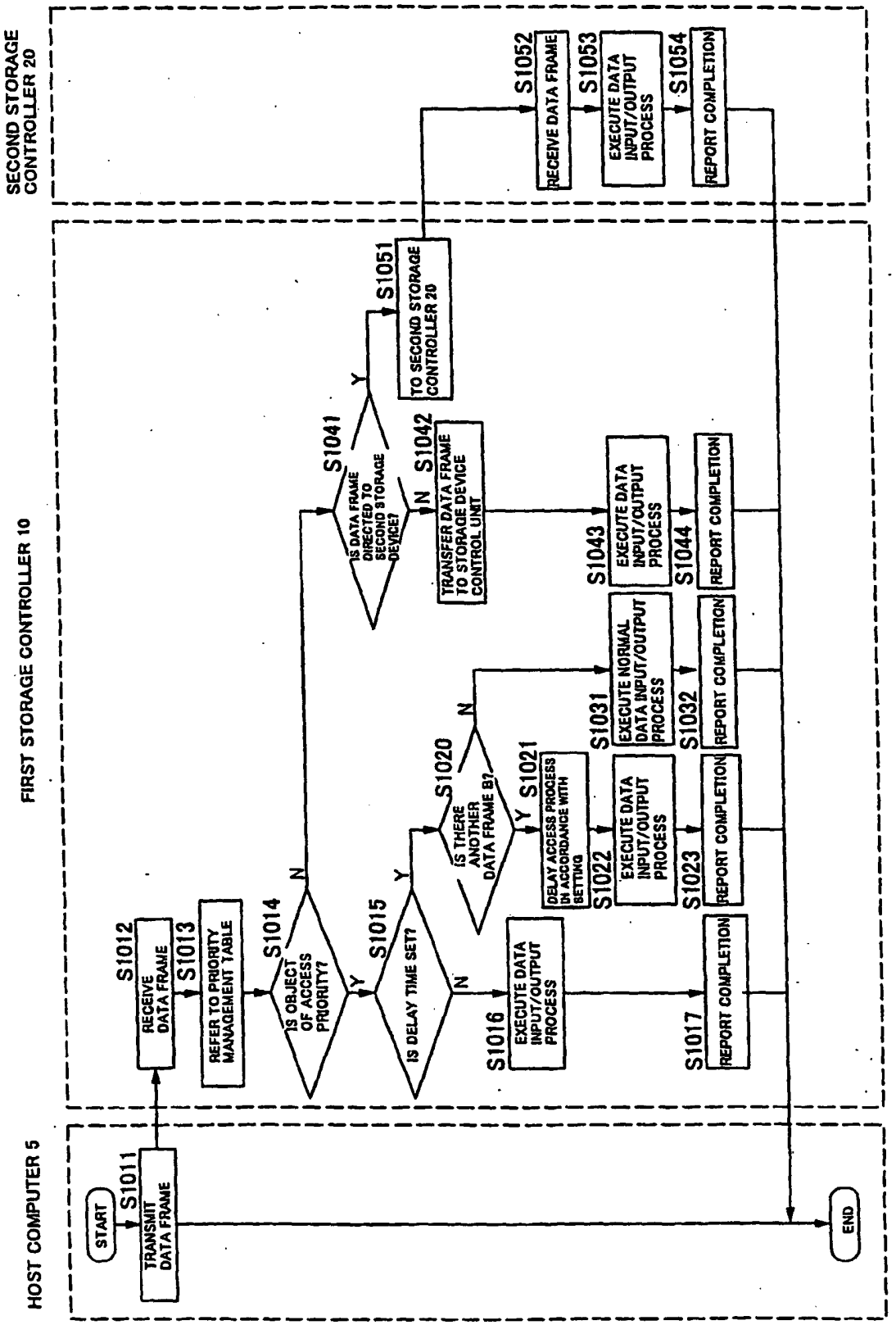

HOST COMPUTER 5
FIRST STORAGE CONTROLLER 10
SECOND STORAGE CONTROLLER 20
START
S1011
TRANSMIT DATA FRAME
S1012
RECEIVE DATA FRAME
S1013
REFER TO PRIORITY MANAGEMENT TABLE
S1014
IS OBJECT OF ACCESS PRIORITY?
S1015
IS DELAY TIME SET?
S1016
EXECUTE DATA INPUT/OUTPUT PROCESS
S1017
REPORT COMPLETION
S1020
IS THERE ANOTHER DATA FRAME B?
S1021
DELAY ACCESS PROCESS IN ACCORDANCE WITH SETTING
S1022
EXECUTE DATA INPUT/OUTPUT PROCESS
S1023
REPORT COMPLETION
S1031
EXECUTE NORMAL DATA INPUT/OUTPUT PROCESS
S1032
REPORT COMPLETION
S1041
IS DATA FRAME DIRECTED TO SECOND STORAGE DEVICE?
S1042
TRANSFER DATA FRAME TO STORAGE DEVICE CONTROL UNIT
S1043
EXECUTE DATA INPUT/OUTPUT PROCESS
S1044
REPORT COMPLETION
S1051
TO SECOND STORAGE CONTROLLER 20
S1052
RECEIVE DATA FRAME
S1053
EXECUTE DATA INPUT/OUTPUT PROCESS
S1054
REPORT COMPLETION
Y
N
END

FIG. 10

FIG. 11

PAIR MANAGEMENT TABLE

LUN(MAIN)	LUN(SUB)
0001	0003
0002	0004
•	•
•	•

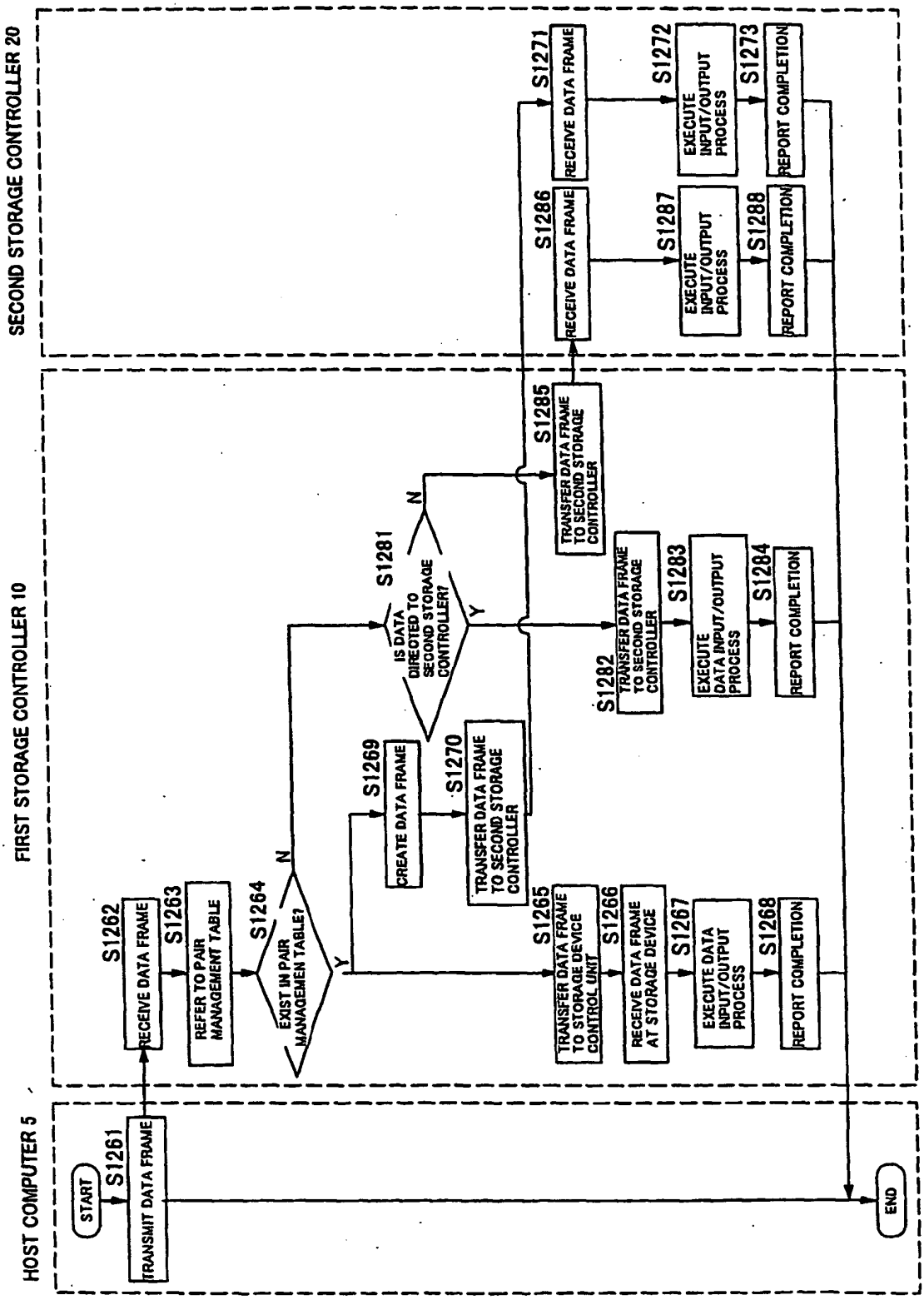

HOST COMPUTER 5
FIRST STORAGE CONTROLLER 10
SECOND STORAGE CONTROLLER 20
START
S1261
TRANSMIT DATA FRAME
S1262
RECEIVE DATA FRAME
S1263
REFER TO PAIR MANAGEMENT TABLE
S1264
EXIST IN PAIR MANAGEMEN TABLE?
N
Y
S1265
TRANSFER DATA FRAME TO STORAGE DEVICE CONTROL UNIT
S1266
RECEIVE DATA FRAME AT STORAGE DEVICE
S1267
EXECUTE DATA INPUT/OUTPUT PROCESS
S1268
REPORT COMPLETION
S1269
CREATE DATA FRAME
S1270
TRANSFER DATA FRAME TO SECOND STORAGE CONTROLLER
S1281
IS DATA DIRECTED TO SECOND STORAGE CONTROLLER?
N
Y
S1282
TRANSFER DATA FRAME TO SECOND STORAGE CONTROLLER
S1283
EXECUTE DATA INPUT/OUTPUT PROCESS
S1284
REPORT COMPLETION
S1285
TRANSFER DATA FRAME TO SECOND STORAGE CONTROLLER
S1271
RECEIVE DATA FRAME
S1272
EXECUTE INPUT/OUTPUT PROCESS
S1273
REPORT COMPLETION
S1286
RECEIVE DATA FRAME
S1287
EXECUTE INPUT/OUTPUT PROCESS
S1288
REPORT COMPLETION
END

FIG. 12

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 20020087751 A1 **[0005]**